

**VOL. XI ISSUE III, Summer (SEPTEMBER-2026)**

# GSSR

## GLOBAL SOCIAL SCIENCES REVIEW

GLOBAL SOCIAL SCIENCES REVIEW (GSSR)

**Title: Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures**

**Abstract**

Existing studies on loyalty to social media influencers have mainly focused on cognitive dissonance, parasocial bonding, the halo effect, social identity, and just-world belief, each in isolation. This paper advances an argument: that these five mechanisms occur not independently but in a sequence, each reducing psychological resistance to the next and producing loyalties more resistant than any single mechanism alone. We call this the compounding loyalty model. Cognitive dissonance is the initial motivation that causes a follower to realize something is wrong; parasocial bonding makes rationalization a requirement, not an option; the halo effect provides a surface-level basis for maintaining trust despite misbehavior; social identity makes loyalty costly to lose; and just-world belief provides a motivated argument that deception did not happen at all. We draw this model from cases including James Charles(2021)and Ramon Abbas, convicted of fraud in 2022. We discuss media literacy, ethical marketing, and platform policy.

**Keywords:** Influencers, Social Media, Cognitive Dissonance, Para-Social Relationships, Halo Effect, Social Identity Theory, Just World Hypothesis, Compounding Loyalty Model

**Author:**

Asma Alolyan: (Corresponding Author)

PhD Psychology, Manchester Metropolitan University,  
Manchester, United Kingdom (UK).  
(Email: [dr.asmaaa3@gmail.com](mailto:dr.asmaaa3@gmail.com))

Pages: 29-41

DOI: 10.31703/gssr.2026(XI-III).03

DOI link: [https://dx.doi.org/10.31703/gssr.2026\(XI-III\).03](https://dx.doi.org/10.31703/gssr.2026(XI-III).03)

Article link: <https://gssrjournal.com/article/why-we-cling-to-flawed-influencers-psychological-mechanisms-behind-loyalty-to-deceptive-social-media-figures>

Full-text Link: <https://gssrjournal.com/article/why-we-cling-to-flawed-influencers-psychological-mechanisms-behind-loyalty-to-deceptive-social-media-figures>

Pdf link: <https://www.gssrjournal.com/jadmin/Author/31rvl0lA2.pdf>

**Global Social Sciences Review**

p-ISSN: [2520-0348](https://doi.org/10.31703/gssr) e-ISSN: [2616-793x](https://doi.org/10.31703/gssr)

DOI(journal):10.31703/gssr

Volume: XI (2026)

DOI (volume):10.31703/gssr.2026(XI)

Issue: III Summer (Sept-2026)

DOI(Issue):10.31703/gssr.2026(XI-III)

Home Page

[www.gssrjournal.com](http://www.gssrjournal.com)

Volume: XI (2026)

<https://www.gssrjournal.com/Current-issue>

Issue: III-Summer (Sept 2026)

<https://www.gssrjournal.com/issue/11/3/2026>

Scope

<https://www.gssrjournal.com/about-us/scope>

Submission

<https://humaglobe.com/index.php/gssr/submissions>

Scan the QR to visit us



Google  
scholar



### Citing this Article

|                |                                                                                                               |
|----------------|---------------------------------------------------------------------------------------------------------------|
| Article Serial | 03                                                                                                            |
| Article Title  | Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures |
| Authors        | Asma Alolyan                                                                                                  |
| DOI            | 10.31703/gssr.2026(XI-III).03                                                                                 |
| Pages          | 29-41                                                                                                         |
| Year           | 2026                                                                                                          |
| Volume         | XI                                                                                                            |
| Issue          | III                                                                                                           |

### Referencing & Citing Styles

|          |                                                                                                                                                                                                                                                                                                                 |
|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| APA      | Alolyan, A. (2026). Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures. <i>Global Social Sciences Review</i> , 11(3), 29-41. <a href="https://doi.org/10.31703/gssr.2026(XI-III).03">https://doi.org/10.31703/gssr.2026(XI-III).03</a>               |
| CHICAGO  | Alolyan, Asma. 2026. "Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures." <i>Global Social Sciences Review</i> 11 (3):29-41. doi: 10.31703/gssr.2026(XI-III).03.                                                                                    |
| HARVARD  | ALOLYAN, A. 2026. Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures. <i>Global Social Sciences Review</i> , 11, 29-41.                                                                                                                              |
| MHRA     | Alolyan, Asma. 2026. 'Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures', <i>Global Social Sciences Review</i> , 11: 29-41.                                                                                                                         |
| MLA      | Alolyan, Asma. "Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures." <i>Global Social Sciences Review</i> 11.3 (2026): 29-41. Print.                                                                                                                 |
| OXFORD   | Alolyan, Asma (2026), 'Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures', <i>Global Social Sciences Review</i> , 11 (3), 29-41.                                                                                                                    |
| TURABIAN | Alolyan, Asma. "Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures." <i>Global Social Sciences Review</i> 11, no. 3 (2026): 29-41. <a href="https://dx.doi.org/10.31703/gssr.2026(XI-III).03">https://dx.doi.org/10.31703/gssr.2026(XI-III).03</a> . |

## Why We Cling to Flawed Influencers: Psychological Mechanisms Behind Loyalty to Deceptive Social Media Figures



Asma Alolyan (Corresponding Author)<sup>1</sup>

<sup>1</sup> PhD Psychology, Manchester Metropolitan University, Manchester, United Kingdom (UK).  
(Email: [dr.asmaa3@gmail.com](mailto:dr.asmaa3@gmail.com))

### Abstract

Existing studies on loyalty to social media influencers have mainly focused on cognitive dissonance, parasocial bonding, the halo effect, social identity, and just-world belief, each in isolation. This paper advances an argument: that these five mechanisms occur not independently but in a sequence, each reducing psychological resistance to the next and producing loyalties more resistant than any single mechanism alone. We call this the compounding loyalty model. Cognitive dissonance is the initial motivation that causes a follower to realize something is wrong; parasocial bonding makes rationalization a requirement, not an option; the halo effect provides a surface-level basis for maintaining trust despite misbehavior; social identity makes loyalty costly to lose; and just-world belief provides a motivated argument that deception did not happen at all. We draw this model from cases including James Charles(2021)and Ramon Abbas, convicted of fraud in 2022. We discuss media literacy, ethical marketing, and platform policy.

**Keywords:** *Influencers, Social Media, Cognitive Dissonance, Para-Social Relationships, Halo Effect, Social Identity Theory, Just World Hypothesis, Compounding Loyalty Model.*

### Introduction:

These social media sites have created a whole new type of “public face,” the influencers who have become popular for their seeming authenticity, relatability, and success. The numbers market products and lifestyles, from skincare to monetary possibilities, and sometimes with hidden business incentives and inaccurate results. The author's case in point is that some years ago after being convinced that a certain fitness influencer's promotional material about smoothies was correct, the author bought a £50 smoothie packet that was found to be just another smoothie made from a repackaged match powder, which is a lot cheaper. This anecdotal experience is echoed by a larger trend: audiences are more likely to stick with influencers even when they discover that this person is being deceptive, sponsored, is making false claims, or is outright lying. This paper investigates the psychological processes underlying that loyalty.

The existing literature provides a good amount of insight into why each of these processes may maintain influencer loyalty, but they are presented as standalone explanations. Studies of parasocial interaction (Sokolova & Kefi, 2020; Sokolova & Perez, 2021) test that mechanism largely on its own; studies of disclosure and persuasion knowledge (Eisend, van Reijmersdal, Boerman, & Tarrahi, 2020) examine the halo effect and source credibility largely in isolation from identity-based motives; and social identity accounts of fan defense (Jenkins, 1992) are rarely connected back to the dissonance-reduction literature that explains why the initial admiration was resistant to revision in the first place. What is missing is an account of how these mechanisms relate to one another, whether they are simply redundant alternative explanations for the same outcome, or whether they interact in a way that makes the resulting loyalty more resistant to correction than any one mechanism could produce alone.



## The Compounding Loyalty Model

This paper argues for the second possibility. We propose that cognitive dissonance, parasocial bonding, the halo effect, social identity, and just-world reasoning do not operate as parallel, redundant explanations but as a sequence, in which each mechanism lowers the resistance the next mechanism subsequently encounters. We term this the compounding loyalty model. In short: Cognitive dissonance is the first mechanism activated, usually in response to encountering disconfirming evidence, and creates a motivated need to rationalize, not revise. The rationalization is not enough to make it seem necessary as opposed to convenient, however, because there is a relational cost to giving up on the admiration of someone felt as a quasi-field, which would not apply to giving up on the admiration of a stranger.

The halo effect then provides an "excuse" for the rationalization that is already going on; attractiveness or polish is interpreted as proof that the disconfirming evidence must be "unreal" or "excessive. Social identity is about turning something private and individual into something public and social, when a follower criticizes the influencer, he or she is not only confronting their beliefs, he or she is confronting him or her in a group. Lastly, any lingering doubts are swept away by providing a general purpose explanation for why the deception, if it exists, will be 'rewarded' sooner or later, allowing the follower to avoid specific action. This type of sequencing creates predictions that simple listing of independent mechanisms doesn't. However, if the mechanisms are truly sequential, as opposed to simply cooccurring, then an intervention that only targets the mechanism activated first in the sequence (e.g., a fact-check aiming at correcting the dissonance inducing claim) should have measurably weaker effects than an intervention that targets both a mechanism earlier in the sequence, and another later in the sequence (e.g., a fact-check aimed at correcting the dissonance inducing claim while also targeting a later stage mechanism like social identity). A prediction of this sort is used directly in section 10.1, and is discussed experimentally in section 11. The rest of the paper elaborates on each mechanism in turn, pointing to the cases and evidence that have justified its inclusion in the series, and then in Section 9 and 10 revisits the compounding strategy, as opposed to the "mechanism by mechanism" approach that has

dominated the literature so far.

30 of 41

The central research question for this analysis is thus double: first, what psychological mechanisms help followers stick with a deceptive influencer once they've come to know it's deceptive and second, are these mechanisms independent and redundant or are they part of a sequential process through which earlier mechanisms can be more effective? Together with empirical studies by Chou, Gaysynsky and Cappella (2020) and Sokolova and Perez (2021), and by relying on documented cases such as that of James Charles and the fraud case of Ramon Abbas (Hushpuppi), we claim the second answer, and uncover its implications for media literacy and consumer behaviour.

## Cognitive Dissonance

Cognitive dissonance is the feeling of discomfort that occurs when a person has two conflicting thoughts, such as liking a particular influencer and knowing that the influencer may be lying (Festinger, 1957). People often develop rationalizations to alleviate this discomfort, instead of changing their feelings of admiration. After all, many of the followers of beauty blogger, James Charles, who has been criticized by the public for promoting scientifically unproven skin care products in 2021, downplayed the controversy and considered exaggeration a normal part of business promotion. Aronson (1992) reports that this type of rationalization is typical when the figure being defended is a figure who has strong emotional investment for the individual, so that the follower is able to maintain his or her initial evaluation even though the evidence is more likely to be a different evaluation. This rationalization is the first in the chain of events that make up the compounding model developed here; by itself it is not definitive and is quickly undermined by other mechanisms that are described in the following paragraphs.

## Parasocial Relationships

Influencers create a feeling of camaraderie like a true friend. In the form of unscripted stories, stand-up videos of daily routines, and videos directly addressed to the viewer, fitness influencer Kayla It Sines

invests in the personal investment of individual viewers. These one-sided emotional bonds are called parasocial relationships and have an effect on the audience's behaviour (Horton & Wohl, [1956](#)). As per Sokolova and Perez ([2021](#)), the researcher noted that people who had a parasocial relationship with the fitness influencers on YouTube were more likely to continue watching the content and for already-watching followers, more likely to exercise, indicating that people who feel a parasocial relationship with the fitness influencer are more likely to continue viewing the content, rather than evaluating it independently. This is reflected in the way people interact with the influencer when he/she is criticized, as seen in the case of Kylie Jenner's criticism after the company's products did not perform as well as they could in 2019. In similar fashion, Sokolova and Kefi ([2020](#)) point out that parasocial connections diminish the critical examination by followers and bring about a sense of being known and trusted by an influencer, transforming him or her into a more intimate than business-like presence.

### **The Halo Effect**

The halo effect is a phenomenon in which people tend to assume that positive attributes such as trustworthiness and competence are attributed to those who are attractive, despite a lack of evidence for such attributes (Thorndike, [1920](#)). For example, Kim Kardashian's SKIMS brand did not face the same problems in terms of commercial performance that other similar lines did in that it was seen as genuine and consumers rated it highly for its performance. (Ohanian, [1990](#)). This bias is also a factor in the comparatively small amount of damage done to the influencer Tana Mongeau after the 2019 TanaCon fiasco, to the extent that a smooth façade can mask problems with handling or even lying to their audience.

### **Social Identity**

Belonging to a particular group or class may be seen as an affiliation to an influencer, and criticism of an influencer may be interpreted as criticism of the follower. The pattern can be explained with the social identity theory (Tajfel & Turner, [1979](#)) that was the result of an online controversy in 2018 for YouTuber Logan Paul, who was criticised for his actions, prompting many of his followers to mobilize to protect his identity, claiming that the situation had been misunderstood. When it comes to defending the criticized person, it appears that group loyalty within fan groups is at the forefront of followers' minds, as Jenkins ([1992](#)) demonstrates.

### **The Just-World Hypothesis**

The just-world hypothesis is that people tend to think that people get what they deserve in life, that success is due to merit, and that evil is ultimately retribution (Lerner, [1980](#)). Before being convicted of fraud in 2022, Ramon Abbas ("Hushpuppi") shared a life of great wealth on

Instagram, with many followers seemingly believing that this was his rightfully earned wealth. To say so would have meant admitting that the social world is not always reliable, that is, that hard work is not always rewarded. As Furnham ([2003](#)) suggests this bias is especially strong when it comes to rich celebrities and can obscure evidence of wrong-doing until it is too obvious to ignore.

### **Mechanisms of Deception**

Whether knowingly or not, influencers engage in the biases outlined above. There are three common trends in the literature:

Exaggerated health claims. The false health based claims about vaccines, treatments, and other relevant topics are rampant on social media and have been found to spread more quickly and farther than the truth; they have been reported to lead to increased vaccine controversy and promotion of untested treatments (Chou, Gaysynsky, & Cappella, [2020](#)). This is a big driver of this trend, especially when it comes to products like 'detox' teas and unproven supplements, which was something that escalated during the COVID -19 pandemic.

Amplification of misinformation. Parasocial trust can amplify false claims; Moran, Swan, and Agajanian (2024) documented how wellness influencers on Instagram used the appearance of relatable, personal storytelling to spread anti vaccine and alternative health misinformation while directly profiting from the products and services they promoted alongside it, as followers extended to these influencers a credibility typically reserved for personal acquaintances.

Promotion of unrealistic standards. Curated and filtered content contributes to body dissatisfaction, particularly among young women, by establishing appearance standards that are difficult or impossible to attain (de Valle, Gallego-García, Williamson, & Wade, 2021); high profile figures such as members of the Kardashian family derive substantial commercial benefit from this dynamic.

## Vulnerable Populations

Susceptibility to these mechanisms is not uniform across the population. A substantial and long running research tradition, distinct from the process focused account developed in this paper, addresses this individual differences question directly: the absorption addiction model of celebrity worship (McCutcheon, Lange, & Houran, 2002) proposes that a compromised sense of personal identity leads some individuals to become psychologically absorbed in a favored celebrity's life as a substitute source of identity and fulfilment, with this absorption taking on addictive qualities that drive escalating attachment. Using the resulting Celebrity Attitude Scale, this line of work identifies a trait like disposition toward parasocial over attachment: McCutcheon, Ashe, Houran, and Maltby (2003) find that individuals scoring highest on the most intense, borderline pathological dimension of celebrity worship show measurable deficits in cognitive flexibility and associative learning relative to lower scoring individuals, while Maltby, Houran, and McCutcheon (2003) link this same dimension to poorer psychological wellbeing. Where the compounding loyalty model proposed in this paper describes a sequence of mechanisms that can operate within any follower confronted with disconfirming evidence, the absorption addiction literature instead identifies a stable subset of individuals for whom parasocial attachment is elevated and more resistant to correction from the outset, a trait level vulnerability that would be expected to accelerate entry into the sequence described in Section 1.1 and to deepen engagement with its later stages, particularly parasocial bonding and social identity, without itself constituting an alternative explanation for the sequence. Individuals exhibiting patterns consistent with compulsive social networking use, who derive reinforcement from engagement metrics such as likes and comments, similarly show reduced critical scrutiny of online content (Kuss & Griffiths, 2017). Younger users are of particular concern. Appearance related social media use is associated with body image concerns among young women, and this relationship is mediated by the tendency to compare one's own appearance with the images encountered (Fardouly, Diedrichs, Vartanian, & Halliwell, 2015). In meta-analyses of experimental and longitudinal evidence, de Valle, Gallego-García, Williamson, and Wade (2021) confirm that exposure to appearance ideal social media imagery has a detrimental effect on body image, and that appearance ideal content is more damaging than other appearance related content. Younger audiences are therefore doubly exposed: to the idealized standards influencers project, and to the commercial claims attached to them.

## Synthesis: The Compounding Loyalty Model

The preceding sections established that five mechanisms, cognitive dissonance, parasocial bonding, the halo effect, social identity, and just-world reasoning, each independently support influencer loyalty. Prior research has largely stopped at this point, establishing that a given mechanism matters within a single study designed around that mechanism alone (e.g., Sokolova & Kefi, 2020, on parasocial interaction; Eisend et al., 2020, on disclosure and the halo effect). Treated this way, the five mechanisms function as a list of alternative explanations, and the natural next empirical question becomes which one explains the most variance, a question several of the studies cited above implicitly treat as central.

We believe that this perspective underestimates the reality of what is going on. The mechanisms are not alternative explanations of the same result, but rather consecutive steps in one process, and are not additive in their explanatory power. Whereas cognitive dissonance is the starting point of the sequence and

creates the motivation to rationalize, parasocial bonding increases the emotional involvement in rationalizing, the halo effect provides the content, social identity transforms this into a public commitment to rationalizing, and just-world reasoning provides the general purpose justification, requiring no interaction with the specific evidence. Each stage does not just add additional explanatory power to the others, it alters the psychological landscape on which the next stage functions, thereby making the next stage more effective than it would be if it was not in the context of the other stages. What makes this compounding structure different from a merely additive structure is that a single fact checking might have the same effect against dissonance alone as against para-social bonding alone, but it does not have as big an impact as it would if the follower had no other reason to hold the attitude besides dissonance, as it does when the attitude has been converted into an emotional commitment. Each of the five rows in Table 1 represents a summarized description of a mechanism, the underlying cognitive or affective process, and the effect on follower loyalty (as well as an example). The compounding claim of the present description is that these five rows are to be read in an ordered sequence, and not as a flat list.

**Table 1**

*The psychological mechanisms that maintain loyalty to misleading influencers in the order of their appearance in the compounding sequence.*

| Mechanism                         | Cognitive/Emotional Process                                             | Effect on Loyalty                                      | Illustrative Example                                                      |
|-----------------------------------|-------------------------------------------------------------------------|--------------------------------------------------------|---------------------------------------------------------------------------|
| Cognitive dissonance              | Rationalizing away evidence of wrongdoing                               | Loyalty persists despite disconfirming evidence        | Dismissing a staged giveaway as normal marketing                          |
| Parasocial bonding                | Influencer is experienced as a personal relationship                    | Followers defend the influencer as they would a friend | Continued purchase of a fitness influencer's program                      |
| Halo effect / attractiveness bias | Physical or aesthetic appeal is inferred as trustworthiness             | Dubious claims are overlooked                          | Uncritical acceptance of a celebrity endorsed product                     |
| Social identity                   | Criticism of the influencer is experienced as an attack on the in group | Fans mobilize to defend the influencer publicly        | Coordinated fan defense of a YouTuber on social media                     |
| Just-world hypothesis             | Success is assumed to be earned                                         | Evidence of fraud is minimized or ignored              | Attributing a public figure's wealth to merit prior to a fraud conviction |
| Individual vulnerability factors  | Validation seeking or compulsive engagement patterns                    | Intensified attachment to influencer figures           | Adolescent users' heightened engagement with idealized content            |

This account makes a prediction that is specific and falsifiable one that is different from a mere list of independent mechanisms: If the compounding model is correct, a particular intervention that targets an early stage mechanism, for example, a fact check aimed at cognitive dissonance, should be less effective the more a given case has already been processed through the sequence of mechanisms, that is, the stronger the parasocial bonding and social identity have already been activated. In contrast, a purely additive model makes no prediction of such interaction; the contribution each mechanism can make to loyalty should be independent and constant throughout the sequence. The experimental design outlined in Section 11 is designed to resolve these two accounts head on.

## Discussion

This paper explores the phenomenon of following and how even the first author, as a follower, is still following influencers despite the evidence of their deception. The results suggest a complex cluster of mental and emotional distortions rather than one single explanation as to how the loyalty is maintained despite clear disconfirming evidence.

### The Compounding Sequence in Practice

In Section 1.1 we proposed that the five mechanisms operate in succession, each reducing the resistance that the next encounters this was rearranged and formalised as a falsifiable claim in Section 10.1, separate from the list of independent biases. We have documented some of the cases throughout this paper and these are intended to be walked through in sequence and not as a catalogue. The sequence begins with cognitive dissonance (Festinger, 1957), when the followers of James Charles were faced with the 2021 skincare controversy, the discomfort was solved by rationalization as the reappraisal of the situation could have been seen as a loss of personal identity, but this rationalization, by itself, was not strong enough and it was just one well documented case of deception. What prevented it from falling apart was the second stage, parasocial attachment (Horton & Wohl, 1956). As fitness influencers are not actual people, but rather props and props that perform a certain role, it becomes easier for them to substitute actual personal closeness with a pattern that can be documented by Sokolova and Perez (2021), where parasocial closeness towards fitness influencers facilitates continued engagement regardless of the accuracy of the fitness influencer's claims, and the pattern is also more emotionally defended because followers are willing to forgive the fitness influencer for replacing actual personal closeness with patterned. The halo effect (Thorndike 1920) then provides that defended rationalization with content and attractiveness is interpreted as evidence that the disconfirming claim must have been an aberration, as Kim Kardashian's ongoing commercial success despite controversy (Ohanian 1990) demonstrates. Social identity processes (Tajfel and Turner, 1979) shift this rationalization from the private to the public sphere, and embeds followers within influencers like Logan Paul, making it harder for them to turn against the influencer than against their "fans," driving up the social cost of doing so beyond the follower's private cost benefit analysis. Finally, the just world hypothesis (Lerner, 1980) completes the round: the need to challenge the specifics of the evidence of Hushpuppi's fraud is not required of followers, just that to assume, in a general and low effort way, that his visible wealth must have been earned, and that anything wrong with it must be discovered and sanctioned by the systems in place would be sufficient. These five cases do not represent five different reasons for loyalty, but instead are five stages in a process which makes the next mechanism that has to work a little more difficult. The continual and interactive nature of social media further increases this structure: a traditional celebrity endorsement would only involve a few direct interactions, whereas influencer marketing is characterised by ongoing parasocial bonding, which is supposed to increase the flow of emotions at the exact moment, which is supposed to be challenging.

34 of 41

### Theoretical and Empirical Alignment

These findings extend classical celebrity endorsement research to the distinctive dynamics of social media, and clarify why that literature's mechanism by mechanism structure has not previously produced a unified account. Erdogan (1999) established that celebrity endorsement leverages perceived credibility and attractiveness; influencer marketing intensifies this effect through direct, ongoing personal engagement. Sokolova and Kefi (2020) found that perceived authenticity and parasocial interaction are key drivers of purchase intention among Instagram and YouTube followers, even when underlying claims are unsubstantiated, a finding our model reframes as evidence for the second stage of the sequence rather than a standalone effect, as in the 2019 TanaCon controversy, where the halo effect (stage three) plausibly built directly on preexisting parasocial trust rather than operating independently of it. Chou, Gaysynsky, and Cappella (2020) further argue that health misinformation circulating on social media exploits both cognitive dissonance and just-world reasoning to sustain audience engagement, a pairing of an early stage and a late stage mechanism that the compounding model predicts should cooccur, since the two are not substitutes for each other but successive links in the same chain. Moran, Swan, and Agajanian (2024) demonstrate that wellness influencers' parasocial closeness with followers facilitated the spread of anti-

vaccine and alternative health misinformation for direct profit, consistent with parasocial bonding's proposed role in converting an initially fragile rationalization into a durable one. Taken together, this literature is more often cited as separate demonstrations of separate biases; the compounding model proposed here offers a way to read it instead as converging evidence for stages of a single underlying process, producing a self-reinforcing cycle in which loyalty sustains engagement and engagement, in turn, deepens loyalty.

### **Implications for Media Literacy, Marketing Ethics, and Policy**

These findings carry several practical implications. First, media literacy interventions should move beyond generic critical thinking instruction to address the specific psychological mechanisms identified here. Given evidence that the appearance of ideal social media content has a measurable detrimental effect on body image, particularly among young women (de Valle et al., [2021](#)), targeted programming, for example, exercises that help users distinguish parasocial closeness from genuine relationships, may be more effective than general literacy instruction alone. Structured exercises simulating the identification of undisclosed sponsorships could help counter cognitive dissonance driven rationalization by normalizing fact-checking as a routine practice, and should also aim to strengthen users' confidence in their own independent judgment.

Second, these dynamics have ramifications for marketing ethics. The use of parasocial trust and related biases could provide a quick boost to the bottom line but would come with long-term reputational and regulatory risks. While the guidelines for endorsing sponsored content by the U.S. Federal Trade Commission (FTC) ([2023](#)) clearly state that disclosure is required, there is still poor adherence to this rule; brands and platforms should better enforce disclosure requirements. Third, these findings bear on platform governance. Policymakers may want to make a more consistent effort to identify unverified claims, similar to the “community notes” style features already available on platforms. Eisend, van Reijmersdal, Boerman and Tarrahi ([2020](#)) conducted a meta-analysis of sponsorship disclosure studies and reported that disclosure is a highly effective way of increasing the recognition of persuasive intent, but the impact on trust and subsequent purchase behaviour is less consistent and depends upon the timing, modality and source of disclosure, meaning that disclosure related policy measures can be effective, but are not necessarily enough. On a more social level, unrealistic body standards set by influencers (de Valle et al., 2021), due to their parasocial trust, are linked to mental health issues, specifically for young women; health misinformation shared by influencers (Moran, Swan, & Agajanian, [2024](#)) has implications for public health outcomes more broadly. Cultural change, alongside regulatory, will likely be necessary to solve these ramifications, with more critical engagement with influencer content becoming the norm.

### **Limitations**

This paper has several limitations. It relies on secondary sources and existing empirical literature rather than original data collection; while this approach allows the argument to be grounded in established findings, it does not provide the granularity that primary research, for example, a survey of platform specific loyalty patterns, would offer. The reflexive account included to illustrate the analysis reflects a single individual's experience and cannot be assumed to generalise; many social media users have not experienced comparable episodes. The evidence base is also predominantly drawn from Western contexts (e.g., James Charles, Hushpuppi); Han and Balabanis's ([2024](#)) meta-analysis of social media influencer research notes that cultural and platform-specific moderators of influencer effectiveness remain under examined, so the extent to which the loyalty dynamics described here generalize to markets such as China or South Korea is currently an open question rather than an established finding. Finally, gender specific patterns beyond body image remain under explored; for example, it is unclear whether men are differentially susceptible to just-world reasoning in the context of financial or cryptocurrency related influencers relative to women. These limitations constrain the generalizability of the present analysis and point to clear directions for future research.

## Future Research

The present analysis is synthetic rather than empirical, and several lines of primary research would substantially strengthen, qualify, or falsify its conclusions. The most direct priority is a test of the compounding loyalty model itself; we outline that design first, followed by four further strands.

Testing the compounding model directly. Section 9 derives a specific prediction: an intervention targeting an early stage mechanism (cognitive dissonance) should lose effectiveness as later stage mechanisms (parasocial bonding, social identity) become more strongly engaged, whereas a purely additive account predicts no such interaction. This could be tested with a between subjects design that first measures or manipulates participants' existing parasocial attachment and identity fusion with a given influencer (low vs. high, using established scales), then exposes all participants to an identical dissonance inducing correction (a documented instance of the influencer's deception), and measures post correction belief revision as the outcome. The compounding model predicts that there will be a significant interaction, correction effectiveness should be significantly less in the high attachment condition, and a strictly additive model predicts there will be a main effect for the correction but no interaction. A secondary design could test the sequence order itself, by experimentally comparing the effects of disrupting the latter stage of the mechanism (e.g., a correction and an in-group identity appeal) to disrupting an early stage mechanism (e.g., a correction) to determine whether disrupting a later stage of the mechanism produces larger gains than repeating the early stage of the mechanism, as the model predicts.

Experimental designs for disclosure and warning interventions. Here are the mechanisms identified, which generate further directly testable hypotheses beyond the compounding prediction above. The presence of the cognitive dissonance effect should lead to rationalization of influencer misbehaviour, and explicit warning labels or regulatory status disclaimers on sponsored content should measurably lessen that rationalization, which could be controlled in a between subjects design with a warning label versus an unlabelled label. In a pair of pre-registered experiments, Bachmann and Valenzuela (2023) demonstrated that, while fact-checks significantly reduced misperceptions, they also significantly decreased people's trust in the news media and their perceptions of media bias, which has not yet been explored in influencer contexts. This makes this an interesting testable question about influencer marketing: Do corrections to influencer claims come at the cost of a more generalized loss of trust, or does repeated exposure to them lead to habituation? Similar designs might include controlling for the content of the message while varying its attractiveness or apparent production quality in order to judge the respective role of the perceived attractiveness of the message and the parasocial closeness.

Longitudinal and scandal response designs. The present account cannot comment on change of loyalty over time, or in the face of repeated evidence of wrongdoing, as the present account is based on cross sectional case examples. Panel studies examining changes in individual followers' parasocial attachment and identity fusion over time before and after a scandal would enable researchers examine whether attachment to the celebrity wanes over time, drops precipitously and then rises back up or remains relatively unchanged except for a smaller portion of the audience. The design that would be very informative and informative would be if you followed the twists and turns of a second, independent scandal for an influencer, so that you could directly test the hypothesis that pre rationalization will reduce or increase the bar set for eventual "disengagement".

Cross cultural replication. The sources of the empirical literature summarized here are largely from Western and mainly Anglophone locations. Han and Balabanis's (2024) meta-analysis of influencer research reveals source credibility and parasocial interaction as the two most prominent explanatory frameworks in the domain, respectively, along with a range of other moderators that are significantly under researched in the field of influencer research, such as cultural and platform specific factors. A social identity driven notion of defence and a lack of responsiveness to individualised fact-checking interventions is likely to be greater among followers of influencers from more collectivist cultural contexts (e.g., China, South Korea, or Southeast Asian countries) than among followers from more individualist cultural contexts (e.g., the United States or the United Kingdom), which leads to the following comparative hypothesis: followers from more collectivist cultural contexts are more likely to exhibit social identity driven defence

and less likely to be responsive to individualized fact-checking interventions when compared with followers from more individualist cultural contexts. It would be difficult to test because there would need to be two sets of data one for WeChat or Xiaohongshu, and another for Instagram and TikTok, each in a similarly distinct cultural and regulatory landscape, and each measured by comparable loyalty and disengagement metrics.

Individual differences and computational approaches. Considering the elements that would predict resistance to disconfirming evidence could be further narrowed down through survey and experimental studies, building on the earlier research on appearance comparison tendency (Fardouly et al., 2015) and into the field of loyalty & disengagement. Separately, the growth of large-scale platform data creates an opportunity for computational approaches: analyses of engagement metrics, comment sentiment, and follower retention curves before and after documented controversies could test at scale whether the loyalty patterns described qualitatively here are visible in behavioural data, and whether they vary systematically by influencer niche (beauty, fitness, finance) or audience demographic.

Neurocognitive mechanisms. Finally, integrating neuroscientific methods, for instance functional imaging of reward related activity during influencer engagement, could help clarify whether the reinforcement patterns implicated in social media use more broadly (Kuss & Griffiths, 2017) share a common neural basis with the loyalty effects described here, or whether parasocial attachment recruits distinct affective circuitry. This work would be more exploratory than the designs above, but could ultimately help distinguish loyalty that is primarily belief based (sustained by rationalization) from loyalty that is primarily reward based (sustained by habitual engagement), a distinction with different implications for intervention design.

## Conclusion

The aim of this paper was to investigate the question of why followers persist in defending influencers who have been revealed to be deceitful, and not stop after the evidence has been brought to light. The compounding loyalty model is the answer that emerges here: None of the other models discussed would explain all of the behaviours observed in the study alone, and the five models analysed are not all alternative explanations for each other but are rather a series of stages, with each one reducing the resistance faced by the next one.

Cognitive dissonance provides the motivation to rationalize; parasocial bonding provides the emotional motivation, and the need to rationalize, to continue to trust; the halo effect provides a superficial basis to continue to trust; social identity provides a social motivation to accept that personal loyalty leads to group loyalty, because the cost of breaking loyalty is higher than it is to the group; and the just-world hypothesis provides a motivation to doubt that the deception ever took place. It is not any one of these biases, but rather this sequence, each stage constraining the next mechanism required to meet the need, that yields the durability seen in cases like James Charles and Hushpuppi, not any list of contributing biases.

This has a practical corollary, and one of the compounding framings makes sharper than a simple list would. Interventions that affect a single mechanism, a single fact-check or a single disclosure label will be only partially effective not just because other mechanisms can be left to take up the belief by the time the intervention comes, but because these other mechanisms will have absorbed and re-anchored the belief by then, which is just the interaction effect that was predicted in Section 9 and specified for testing in Section 11. The implication for practice is that effective countermeasures will require multiple and sequential strategies strategies that include disclosure, as well as strategies for fact checking and strategies for media literacy that explicitly identify Para social attachment and in-group loyalty as psychological phenomena, not simply act as generic critical thinking skills that will "transfer" to content from influencers.

More bodily, the persistence of this loyalty is not best understood as a failure of individual judgement. The mechanisms reviewed here dissonance reduction, attachment, identity protection, and belief in a just world are commonplace mechanisms of cognition with adaptive functions in other domains of social interaction, and they are merely provided with a particularly high frequency and a particularly personalized array of opportunities in social media environments that unfold in opposition to the interests of their

followers. Considering the issue of influencer loyalty in such terms as the inevitable byproduct of familiar psychological dynamics in a new setting, not as duplicity could prove a valuable corrective as well: It shifts the focus from "critical thinking" to the platform design and regulatory system, while leaving duplicity untouched; it also provides researchers, educators, and policymakers with a more manageable set of levers, and a testable model to test them against, than it does with "critical thinking" alone.

## References

- Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry*, 3(4), 303–311. [https://doi.org/10.1207/s15327965pli0304\\_1](https://doi.org/10.1207/s15327965pli0304_1)  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Bachmann, I., & Valenzuela, S. (2023). Studying the downstream effects of fact-checking on social media: Experiments on correction formats, belief accuracy, and media trust. *Social Media + Society*, 9(2), 1–13. <https://doi.org/10.1177/20563051231179694>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Chou, W.-Y. S., Gaysynsky, A., & Cappella, J. N. (2020). Where we go from here: Health misinformation on social media. *American Journal of Public Health*, 110(S3), S273–S275. <https://doi.org/10.2105/AJPH.2020.305905>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- de Valle, M. K., Gallego-García, M., Williamson, P., & Wade, T. D. (2021). Social media, body image, and the question of causation: Meta-analyses of experimental and longitudinal evidence. *Body Image*, 39, 276–292. <https://doi.org/10.1016/j.bodyim.2021.10.001>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Eisend, M., van Reijmersdal, E. A., Boerman, S. C., & Tarrahi, F. (2020). A meta-analysis of the effects of disclosing sponsored content. *Journal of Advertising*, 49(3), 344–366. <https://doi.org/10.1080/00913367.2020.1765909>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Erdogan, B. Z. (1999). Celebrity endorsement: A literature review. *Journal of Marketing Management*, 15(4), 291–314. <https://doi.org/10.1362/026725799784870379>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Fardouly, J., Diedrichs, P. C., Vartanian, L. R., & Halliwell, E. (2015). Social comparisons on social media: The impact of Facebook on young women's body image concerns and mood. *Body Image*, 13, 38–45. <https://doi.org/10.1016/j.bodyim.2014.12.002>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Federal Trade Commission. (2023). FTC's endorsement guides: What people are asking. <https://www.ftc.gov/business-guidance/resources/ftcs-endorsement-guides-what-people-are-asking>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences*, 34(5), 795–817. [https://doi.org/10.1016/S0191-8869\(02\)00072-7](https://doi.org/10.1016/S0191-8869(02)00072-7)  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)
- Han, J., & Balabanis, G. (2024). Meta-analysis of social media influencer impact: Key antecedents and theoretical foundations. *Psychology & Marketing*, 41(2), 394–426. <https://doi.org/10.1002/mar.21927>  
[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, 19(3), 215–229.

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Jenkins, H. (1992). *Textual poachers: Television fans and participatory culture*. Routledge.

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Kuss, D. J., & Griffiths, M. D. (2017). Social networking sites and addiction: Ten lessons learned. *International Journal of Environmental Research and Public Health*, 14(3), 311. <https://doi.org/10.3390/ijerph14030311>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. Plenum Press.

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Maltby, J., Houran, J., & McCutcheon, L. E. (2003). A clinical interpretation of attitudes and behaviors associated with celebrity worship. *Journal of Nervous and Mental Disease*, 191(1), 25–29. <https://doi.org/10.1097/00005053-200301000-00005>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

McCutcheon, L. E., Ashe, D. D., Houran, J., & Maltby, J. (2003). A cognitive profile of individuals who tend to worship celebrities. *The Journal of Psychology*, 137(4), 309–322. <https://doi.org/10.1080/00223980309600616>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

McCutcheon, L. E., Lange, R., & Houran, J. (2002). Conceptualization and measurement of celebrity worship. *British Journal of Psychology*, 93(1), 67–87. <https://doi.org/10.1348/000712602162454>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Moran, R. E., Swan, A. L., & Agajanian, T. (2024). Vaccine misinformation for profit: Conspiratorial wellness influencers and the monetization of alternative health. *International Journal of Communication*, 18, 1202–1224. <https://ijoc.org/index.php/ijoc/article/view/21128>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3), 39–52. <https://doi.org/10.1080/00913367.1990.10673191>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Sokolova, K., & Kefi, H. (2020). Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of Retailing and Consumer Services*, 53, 101742. <https://doi.org/10.1016/j.jretconser.2019.01.011>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Sokolova, K., & Perez, C. (2021). You follow fitness influencers on YouTube. But do you actually exercise? How parasocial relationships, and watching fitness influencers, relate to intentions to exercise. *Journal of Retailing and Consumer Services*, 58, 102276. <https://doi.org/10.1016/j.jretconser.2020.102276>

[Google Scholar](#) [Worldcat](#) [Fulltext](#)

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.

[Google Scholar](#)   [Worldcat](#)   [Fulltext](#)

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4(1), 25–29. <https://doi.org/10.1037/h0071663>

[Google Scholar](#)   [Worldcat](#)   [Fulltext](#)